PERGAMON
Applied Mathematics Letters 15 (2002) 153–157
Applied Mathematics Letters
www.elsevier.com/locate/aml

Probabilistic Inequalities and Remarks

G. A. ANASTASSIOU
Department of Mathematical Sciences
The University of Memphis
Memphis, TN 38152, U.S.A.
ganastss@memphis.edu

V. G. PAPANICOLAOU
Department of Mathematics and Statistics
Wichita State University
Wichita, KS 67260-0033, U.S.A.
papanico@math.twsu.edu

(Received March 2001; accepted May 2001)

Abstract—The authors use their recently proved integral inequality to obtain bounds for the covariance of two random variables

(a) in a general setup and
(b) for a class of special joint distributions.

The same inequality is also used to estimate the difference of the expectations of two random variables. Finally, the authors study the attainability of a related inequality. © 2002 Elsevier Science Ltd. All rights reserved.

Keywords—Covariance, Joint distribution, Multidimensional sharp integral inequality, Partial derivative.

1. INTRODUCTION

In their recent article [1], the authors proved the following result.

THEOREM A. *Let $f \in C^n(B)$, $n \in \mathbb{N}$, where $B = [a_1, b_1] \times \cdots \times [a_n, b_n]$, $a_j, b_j \in \mathbb{R}$, with $a_j < b_j$, $j = 1, \dots, n$. Denote by ∂B the boundary of the box B. Assume that $f(x) = 0$, for all $x = (x_1, \dots, x_n) \in \partial B$ (in other words, we assume that $f(\dots, a_j, \dots) = f(\dots, b_j, \dots) = 0$, for all $j = 1, \dots, n$). Then*

$$\int_B |f(x_1, \dots, x_n)| \, dx_1 \cdots dx_n \le \frac{m(B)}{2^n} \int_B \left| \frac{\partial^n f(x_1, \dots, x_n)}{\partial x_1 \cdots \partial x_n} \right| dx_1 \cdots dx_n, \tag{1}$$

where $m(B) = \prod_{j=1}^n (b_j - a_j)$ is the n^{th}-dimensional volume (i.e., the Lebesgue measure) of B.

In the present article, we give probabilistic applications of the above inequality and related remarks.

PII: S0893-9659(01)00110-0

2. APPLICATIONS

Let $X \geq 0$ be a random variable. Given a sequence of positive numbers $b_n \to \infty$, we introduce the events

$$B_n = \{X > b_n\}.$$

Then (as in the proof of Chebychev's inequality)

$$E\left[X 1_{B_n}\right] = \int_{B_n} X\, dP \geq b_n P(B_n) = b_n P(X > b_n), \tag{2}$$

where, as usual, P is the corresponding probability measure and E its associated expectation. If $E[X] < \infty$, then monotone convergence gives (since $X 1_{B_n^c} \nearrow X$)

$$\lim_n E\left[X 1_{B_n^c}\right] = E[X].$$

Thus,

$$E\left[X 1_{B_n}\right] = E[X] - E\left[X 1_{B_n^c}\right] \to 0.$$

Hence, by (2),

$$b_n P(X > b_n) \to 0, \tag{3}$$

for any sequence $b_n \to \infty$.

The following proposition is well known, but we include it here for the sake of completeness.

PROPOSITION 1. *Let $X \geq 0$ be a random variable with (cumulative) distribution function $F(x)$. Then*

$$E[X] = \int_0^\infty [1 - F(x)]\, dx \tag{4}$$

(notice that if the integral in the right-hand side diverges, it should diverge to $+\infty$, and the equality still makes sense).

PROOF. Integration by parts gives

$$\int_0^\infty [1 - F(x)]\, dx = \lim_{x\to\infty} x[1 - F(x)] + \int_0^\infty x\, dF(x) = \lim_{x\to\infty} xP(X > x) + E[X].$$

If $E[X] < \infty$, then (4) follows by (3). If $E[X] = \infty$, then the above equation implies that

$$\int_0^\infty [1 - F(x)]\, dx = \infty,$$

since $xP(X > x) \geq 0$, whenever $x \geq 0$. ∎

REMARK 1. If $X \geq a$ is a random variable with distribution function $F(x)$, we can apply Proposition 1 to $Y = X - a$ and get

$$E[X] = a + \int_a^\infty [1 - F(x)]\, dx. \tag{5}$$

We now present some applications of Theorem A.

PROPOSITION 2. *Consider two random variables X and Y taking values in the interval $[a, b]$. Let $F_X(x)$ and $F_Y(x)$, respectively, be their distribution functions, which are assumed absolutely continuous. Their corresponding densities are $f_X(x)$ and $f_Y(x)$, respectively. Then*

$$|E[X] - E[Y]| \leq \frac{b-a}{2} \int_a^b |f_X(x) - f_Y(x)|\, dx. \tag{6}$$

PROOF. We have that $F_X(b) = F_Y(b) = 1$, thus formula (5) implies

$$E[X] = a + \int_a^b [1 - F_X(x)]\,dx \qquad \text{and} \qquad E[Y] = a + \int_a^b [1 - F_Y(x)]\,dx.$$

It follows that

$$|E[X] - E[Y]| \le \int_a^b |F_X(x) - F_Y(x)|\,dx.$$

Since $F_X(x)$ and $F_Y(x)$ are absolutely continuous with densities $f_X(x)$ and $f_Y(x)$, respectively, and since $F_X(a) - F_Y(a) = F_X(b) - F_Y(b) = 0$, Theorem A, together with the above inequality, imply (6). ∎

THEOREM 1. *Let X and Y be two random variables with joint distribution function $F(x,y)$. We assume that $a \le X \le b$ and $c \le Y \le d$. We also assume that $F(x,y)$ possesses a (joint) probability density function $f(x,y)$. Then*

$$|\operatorname{cov}(X,Y)| \le \frac{(b-a)(d-c)}{4} \int_c^d \int_a^b |f(x,y) - f_X(x) f_Y(y)|\,dx\,dy, \tag{7}$$

where

$$\operatorname{cov}(X,Y) = E[XY] - E[X]\,E[Y]$$

is the covariance of X and Y, while $f_X(x)$ and $f_Y(y)$ are the marginal densities of X and Y, respectively.

PROOF. Integration by parts gives

$$E[XY] = aE[Y] + cE[X] - ac + \int_c^d \int_a^b [1 - F(x,d) - F(b,y) + F(x,y)]\,dx\,dy \tag{8}$$

(in fact, this formula generalizes to n random variables). Notice that $F(x,d)$ is the (marginal) distribution function of X alone, and similarly, $F(b,y)$ is the (marginal) distribution function of Y alone. Thus, by (5),

$$E[X] = a + \int_a^b [1 - F(x,d)]\,dx \qquad \text{and} \qquad E[Y] = c + \int_c^d [1 - F(b,y)]\,dy,$$

which gives

$$(E[X] - a)\,(E[Y] - c) = \int_c^d \int_a^b [1 - F(x,d)]\,[1 - F(b,y)]\,dx\,dy$$

or

$$E[X]\,E[Y] = aE[Y] + cE[X] - ac + \int_c^d \int_a^b [1 - F(x,b) - F(a,y) + F(x,d)F(b,y)]\,dx\,dy. \tag{9}$$

Subtracting (9) from (8), we obtain

$$\operatorname{cov}(X,Y) = E[XY] - E[X]\,E[Y] = \int_c^d \int_a^b [F(x,y) - F(x,d)F(a,y)]\,dx\,dy.$$

Thus,

$$|\operatorname{cov}(X,Y)| \le \int_c^d \int_a^b |F(x,y) - F(x,d)F(b,y)|\,dx\,dy.$$

Now $F(a,y) = F(x,c) = 0$ and $F(b,d) = 1$. It follows that $F(x,y) - F(x,d)F(b,y)$ vanishes on the boundary of $B = (a,b) \times (c,d)$. Hence Theorem A, together with the above inequality, imply (7). ∎

In Feller's classical book [2, Par. 5, p. 165], we find the following fact. Let $F(x)$ and $G(y)$ be distribution functions on $\mathbb{R}$, and set

$$U(x,y) = F(x)G(y)\left\{1 - \alpha\left[1 - F(x)\right]\left[1 - G(y)\right]\right\}, \tag{10}$$

where $-1 \le \alpha \le 1$. Then $U(x,y)$ is a distribution function on $\mathbb{R}^2$, with marginal distributions $F(x)$ and $G(y)$. Furthermore, $U(x,y)$ possesses a density if and only if $F(x)$ and $G(y)$ do.

THEOREM 2. *Let X and Y be two random variables with joint distribution function $U(x,y)$, as given by (10). For the marginal distributions $F(x)$ and $G(y)$, of X and Y, respectively, we assume that they possess densities $f(x)$ and $g(y)$ (respectively). Furthermore, we assume that $a \le X \le b$ and $c \le Y \le d$. Then*

$$|\mathrm{cov}(X,Y)| \le |\alpha| \frac{(b-a)(d-c)}{16}. \tag{11}$$

PROOF. The (joint) density $u(x,y)$ of $U(x,y)$ is

$$u(x,y) = \frac{\partial^2 U(x,y)}{\partial x \partial y} = f(x)g(y)\left\{1 - \alpha\left[1 - 2F(x)\right]\left[1 - 2G(y)\right]\right\}.$$

Hence,

$$u(x,y) - f(x)g(y) = -\alpha\left[1 - 2F(x)\right]\left[1 - 2G(y)\right] f(x)g(y),$$

and Theorem 1 gives

$$|\mathrm{cov}(X,Y)| \le |\alpha| \frac{(b-a)(d-c)}{4} \left[\int_a^b |1 - 2F(x)|\, f(x)\, dx\right] \left[\int_c^d |1 - 2G(y)|\, g(y)\, dy\right]$$

or

$$|\mathrm{cov}(X,Y)| \le |\alpha| \frac{(b-a)(d-c)}{4} E\left[|1 - 2F(X)|\right] E\left[|1 - 2G(Y)|\right]. \tag{12}$$

Now, $F(x)$ and $G(y)$ have density functions, hence, $F(X)$ and $G(Y)$ are uniformly distributed on $(0,1)$. Thus,

$$E\left[|1 - 2F(X)|\right] = E\left[|1 - 2G(Y)|\right] = \frac{1}{2},$$

and the proof is finished by using the above equalities in (12). ∎

3. REMARKS ON AN INEQUALITY

The basic ingredient in the proof of Theorem A is the following inequality (also shown in [1]).

Let $B = (a_1, b_1) \times \cdots \times (a_n, b_n) \subset \mathbb{R}^n$, with $a_j < b_j$, $j = 1, \dots, n$. Denote by $\bar{B}$ and ∂B, respectively, the (topological) closure and the boundary of the open box B. Consider the functions $f \in C_0(\bar{B}) \cap C^n(B)$, namely, the functions that are continuous on $\bar{B}$, have n continuous derivatives in B, and vanish on ∂B. Then for such functions, we have

$$|f(x_1, \dots, x_n)| \le \frac{1}{2^n} \int_B \left| \frac{\partial^n f(x_1, \dots, x_n)}{\partial x_1 \cdots \partial x_n} \right| dx_1 \cdots dx_n,$$

true for all $(x_1, \dots, x_n) \in B$. In other words,

$$\|f\|_\infty \le \frac{1}{2^n} \int_B \left| \frac{\partial^n f(x_1, \dots, x_n)}{\partial x_1 \cdots \partial x_n} \right| dx_1 \cdots dx_n, \tag{13}$$

where $\|\cdot\|_\infty$ is the supnorm of $C_0(\bar{B})$.

REMARK 1. Suppose we have that $a_j = -\infty$ for some (or all) j and/or $b_j = \infty$ for some (or all) j. Let $\hat{B} = \bar{B} \cup \{\infty\}$ be the one-point compactification of $\bar{B}$ and assume that $f \in C_0(\hat{B}) \cap C^n(B)$. This means that if $a_j = -\infty$, then $\lim_{s_j \to -\infty} f(x_1, \dots, s_j, \dots, x_n) = 0$, for all $x_k \in (a_k, b_k)$, $k \ne j$, and also that if $b_j = \infty$, then $\lim_{s_j \to \infty} f(x_1, \dots, s_j, \dots, x_n) = 0$, for all $x_k \in (a_k, b_k)$, $k \ne j$.

Then the proof of (13), as given in [1], remains valid.

REMARK 2. In the case of an unbounded domain, as described in the previous remark, although $\|f\|_\infty < \infty$, it is possible that

$$\int_B \left| \frac{\partial^n f(x_1, \ldots, x_n)}{\partial x_1 \cdots \partial x_n} \right| dx_1 \cdots dx_n = \infty.$$

For example, take $B = (0, \infty)$ and

$$f(x) = \int_0^x \left(e^{i\xi^2} - \frac{\sqrt{\pi}}{2} e^{i\pi/4} e^{-\xi} \right) d\xi.$$

Before discussing sharpness and attainability for (13), we need some notation. For $c = (c_1, \ldots, c_n) \in B$, we introduce the intervals

$$I_{j,0}(c) = (a_j, c_j) \quad \text{and} \quad I_{j,1}(c) = (c_j, b_j), \qquad j = 1, \ldots, n.$$

THEOREM 3. *For $f \in C_0(\bar{B}) \cap C^n(B)$, inequality (13) is sharp. Equality is attained if and only if there is a $c = (c_1, \ldots, c_n) \in B$ such that the n^{th} mixed derivative $\frac{\partial^n f(x_1, \ldots, x_n)}{\partial x_1 \cdots \partial x_n}$ does not change sign in each of the 2^n "sub-boxes" $I_{1,\varepsilon_1}(c) \times \cdots \times I_{n,\varepsilon_n}(c)$.*

PROOF. ($\Leftarrow$) For an arbitrary $c = (c_1, \ldots, c_n) \in B$, we have

$$f(c_1, \ldots, c_n) = (-1)^{\varepsilon_1 + \cdots + \varepsilon_n} \int_{I_{1,\varepsilon_1}(c)} \cdots \int_{I_{n,\varepsilon_n}(c)} \frac{\partial^n f(x_1, \ldots, x_n)}{\partial x_1 \cdots \partial x_n} dx_1 \cdots dx_n, \tag{14}$$

where each ε_j can be either 0 or 1. If c is as in the statement of the theorem, then

$$|f(c_1, \ldots, c_n)| = \int_{I_{1,\varepsilon_1}(c)} \cdots \int_{I_{n,\varepsilon_n}(c)} \left| \frac{\partial^n f(x_1, \ldots, x_n)}{\partial x_1 \cdots \partial x_n} \right| dx_1 \cdots dx_n. \tag{15}$$

Adding up (15) for all 2^n choices of $(\varepsilon_1, \ldots, \varepsilon_n)$, we obtain

$$2^n |f(c_1, \ldots, c_n)| = \sum_{\varepsilon_1, \ldots, \varepsilon_n} \int_{I_{1,\varepsilon_1}(c)} \cdots \int_{I_{n,\varepsilon_n}(c)} \left| \frac{\partial^n f(x_1, \ldots, x_n)}{\partial x_1 \cdots \partial x_n} \right| dx_1 \cdots dx_n \tag{16}$$

or

$$|f(c_1, \ldots, c_n)| = \frac{1}{2^n} \int_B \left| \frac{\partial^n f(x_1, \ldots, x_n)}{\partial x_1 \cdots \partial x_n} \right| dx_1 \cdots dx_n.$$

This forces (13) to become an equality. In fact, we also must have $|f(c_1, \ldots, c_n)| = \|f\|_\infty$, i.e., $|f(x)|$ attains its maximum at $x = c$.

($\Rightarrow$) Conversely, let $c \in B$ be a point at which $|f(c)| = \|f\|_\infty$ (such a c exists since f is continuous and vanishes on ∂B and at infinity). If for this c, the sign of the mixed derivative $\frac{\partial^n f(x_1, \ldots, x_n)}{\partial x_1 \cdots \partial x_n}$ changes inside some "sub-box" $I_{1,\varepsilon_1}(c) \times \cdots \times I_{n,\varepsilon_n}(c)$, then (14) implies

$$\|f\|_\infty = |f(c_1, \ldots, c_n)| < \int_{I_{1,\varepsilon_1}(c)} \cdots \int_{I_{n,\varepsilon_n}(c)} \left| \frac{\partial^n f(x_1, \ldots, x_n)}{\partial x_1 \cdots \partial x_n} \right| dx_1 \cdots dx_n.$$

Thus, for this f, (13) becomes a strict inequality. ∎

REFERENCES

1. G.A. Anastassiou and V.G. Papanicolaou, A new basic sharp integral inequality, (preprint).
2. W. Feller, *An Introduction to Probability Theory and its Applications, Volume II*, Second Edition, J. Wiley, New York, (1971).